

## INVARIANT SETS FOR QMF FUNCTIONS

ADAM JONSSON

ABSTRACT. A quadrature mirror filter (QMF) function can be considered as a  $g$ -function on a space of binary sequences. The QMF functions that give rise to multiresolution analyses are then distinguished by their invariant sets. By characterizing them, we answer in the affirmative a question raised by Gundy (*Notices Americ. Math. Soc.*, (9)57:1094-1104, 2010). Invariant subshifts for continuous  $g$ -functions are characterized by necessary and sufficient conditions.

## 1. INTRODUCTION

The main question of this paper concerns stationary distributions for a classic Markov process. However, the question arose in fairly recent constructions of wavelet bases. It is the purpose of this section to provide background and give a summary of the main results.

We begin with some notation that will be used throughout. Let  $\{0, 1\}$  have the discrete topology, let  $\mathbb{Z}_- = \{0, -1, -2, \dots\}$ , and let  $\Theta$  be the shift on  $X = \{0, 1\}^{\mathbb{Z}_-}$ . Points of  $X$  are written  $\xi = (\dots, x_{-1}, x_0)$ . Let  $\mathcal{G}$  denote the  $g$ -functions on  $X$ . That is,  $\mathcal{G}$  is the set of all nonnegative functions on  $X$  with the property that

$$\sum_{\xi' \in \Theta^{-1}(\xi)} g(\xi') = 1 \text{ for every } \xi \in X. \quad (1)$$

If  $K \subset X$  is a subshift, meaning that  $K$  is closed and satisfies  $K = \{\Theta(\xi) : \xi \in K\}$ , and if  $g \in \mathcal{G}$ , then we say that  $K$  is  $g$ -invariant or invariant for  $g$  if<sup>1</sup>

$$g(\xi) = 0 \text{ for all } \xi \in K^c \text{ such that } \Theta(\xi) \in K. \quad (2)$$

The notion of a  $g$ -function has its origin in the probability theory of “chains with infinite connections”. Some of the first contributions to this subject were made by Doeblin and Fortet [7]. Keane [15] introduced  $g$ -functions in ergodic theory. Much of the large literature on  $g$ -functions that has evolved since then concerns the problem of finding criteria for uniqueness of associated invariant measures, also called  $g$ -measures. See for example Berbee [1], Johansson and Öberg [13] and Stenflo [21, 20] for details and references. This paper concerns a class of  $g$ -functions that have more than one  $g$ -measure. The support sets of these measures give examples of invariant subshifts. Our motivation for studying them comes from the characterization of scaling functions for multiresolution analyses (MRA). In this context, a necessary condition for a nonnegative, 1-periodic function  $p(\xi), \xi \in \mathbb{R}$ ,

---

*Date:* October 14, 2015.

*Key words and phrases.* invariance  $\times 2$ ; Markov operators;  $g$ -functions; scaling functions.

I would like to thank Dick Gundy for conversations on the topic of this paper, Wolfgang Krieger for information about  $g$ -function presentations, and Jeff Steif for pointing out errors in an earlier version of the manuscript. I am of course responsible for any remaining errors.

<sup>1</sup>We let  $E^c$  and  $\overline{E}$  denote complement and closure, respectively.

to be the squared modulus of a lowpass filter — that is, the Fourier generator of a scaling function for a MRA with respect to dyadic dilation — is the “quadrature mirror filter” (QMF) condition:

$$p(\xi/2) + p(\xi/2 + 1/2) = 1 \text{ for every } \xi \in [0, 1]. \quad (3)$$

Let  $\mathcal{Q}$  be the class of nonnegative 1-periodic functions for which (3) holds and  $p(0) = 1$ . Call them quadrature mirror filter (QMF) functions. Any function on  $[0, 1]$  can be lifted to  $X$  via composition with  $\tau: X \rightarrow [0, 1]$ , where  $\tau(\xi) := \sum_{j=-\infty}^0 x_j 2^{j-1}$ . Since (3) implies that  $p \circ \tau$  satisfies (1),  $p \circ \tau \in \mathcal{G}$  if  $p \in \mathcal{Q}$ .

We will say that  $p \in \mathcal{Q}$  generates a scaling function, writing  $p \in \mathcal{Q}^{\text{MRA}}$ , if the convergent (since  $0 \leq p \leq 1$ ) infinite product

$$\hat{\Phi}_p(\xi) := \prod_{j=1}^{\infty} p(\xi/2^j), \xi \in \mathbb{R}, \quad (4)$$

has the property that

$$(a) \sum_{k \in \mathbb{Z}} \hat{\Phi}_p(\xi + k) = 1 \text{ for a.e } \xi \in [0, 1], \quad (5)$$

$$(b) \lim_{j \rightarrow \infty} \hat{\Phi}_p(2^{-j}\xi) = 1 \text{ for a.e } \xi \in \mathbb{R}. \quad (6)$$

Many authors have studied the problem of finding constraints on  $p$  for  $\hat{\Phi}_p$  to satisfy (a) and (b). The first necessary and sufficient conditions were found for trigonometric polynomial  $p \in \mathcal{Q}$  by Cohen [3] and Lawton [16]. Cohen later showed that his condition (see Section 6) is necessary and sufficient for smooth  $p$ , and his approach was developed further by Hernández and Weiss [12] and by Papadakis, Šikić and Weiss [18], and others. That Cohen’s condition is not necessary if  $p$  is only assumed continuous was discovered by Dobrić et al [6]. They view  $p$  as the transition function for a Markov process  $\xi_0, \xi_1, \dots$  on the unit interval. From a fixed  $\xi_0 \in [0, 1]$ , the process is governed by a probability  $P_p(|\xi_0)$  determined by  $p$  and evolves according to

$$\xi_{t+1} = \xi_t/2 \text{ or } \xi_t/2 + 1/2, \quad (7)$$

with conditional transition-probabilities  $p(\xi_t/2)$  and  $p(\xi_t/2 + 1/2)$ , respectively. A set  $B \subset [0, 1]$  is *invariant* for  $p$  if  $P_p(\xi_1 \in B | \xi_0) = 1$  for all  $\xi_0 \in B$ . Since  $p(0) = 1$  implies  $p(1/2) = 0$ ,  $B := \{0, 1\}$  is invariant for every  $p \in \mathcal{Q}$ . The basic conclusion in [6] says that  $\sum_{k \in \mathbb{Z}} \hat{\Phi}_p(\xi_0 + k)$  gives the  $P_p(|\xi_0)$ -probability that  $\xi_1, \xi_2, \dots$  converges to this invariant set (cf. Section 6). So if  $p$  generates a scaling function, then  $\xi_1, \xi_2, \dots$  must converge to 0 or 1 almost surely for Lebesgue almost every  $\xi_0 \in [0, 1]$ . When  $p$  is so smooth that  $\sum_{k \in \mathbb{Z}} \hat{\Phi}_p(\xi + k)$  is a continuous function of  $\xi$ , the equality in (a) must hold for all  $\xi \in [0, 1]$ . However, if  $p$  is continuous but non-smooth, then the left hand side of this equality may vanish on invariant sets of measure zero.

Such “inaccessible” invariant sets are incompatible with Cohen’s condition. Dobrić et al produce a continuous  $p \in \mathcal{Q}^{\text{MRA}}$  such that  $B := \{1/3, 2/3\}$  is invariant. Gundy [9] studies QMF functions with more general invariant sets. He considers the Markov process on  $\{0, 1\}^{\mathbb{Z}^-}$  with transition function  $p \circ \tau$  to construct a collection of  $p \in \mathcal{Q}^{\text{MRA}}$  with Cantor-type inaccessible sets, then raises the question of whether (a) and (b) place restrictions on invariant sets associated with  $p$  ([10, p.

1103]). Since every closed subset of  $(0, 1)$  invariant under the doubling map

$$\theta(\xi) := 2\xi \pmod{1} \quad (8)$$

has Lebesgue measure zero by the ergodicity of  $\theta$ , we should ask: *Given a closed,  $\theta$ -invariant  $B \subset (0, 1)$ , is there a continuous  $p \in \mathcal{Q}^{MRA}$  for which  $B$  is invariant?*

Our primary result gives a sharp answer to this question. By first establishing that if  $B \subset (0, 1)$  is invariant for a continuous  $p \in \mathcal{Q}$ , then

$$\overline{B_e^c} \cap \{\xi^* : \xi \in \overline{B_e^c}\} = \emptyset, \quad (9)$$

where  $B_e^c = \{\xi \in (0, 1) \cap B^c : \theta(\xi) \in B\}$ ,  $\xi^* = \xi + 1/2 \pmod{1}$ , we show that it is essentially affirmative:

**Theorem 1.** *Every closed,  $\theta$ -invariant  $B \subset (0, 1)$  that satisfies (9) is invariant for some continuous  $p \in \mathcal{Q}^{MRA}$ .*

Passing from QMF functions to probability with  $g$ -functions is not only a convenience. In [11], this approach was used to describe scaling functions for MRA with respect to  $2 \times 2$  dilation matrices of determinant  $\pm 2$ . The second goal of the present paper is to characterize invariant subshifts for continuous  $g$ -functions on  $X = \{0, \dots, J\}^{\mathbb{Z}^-}$ ,  $J \geq 1$ . The case  $J = 1$  is related to Theorem 1. A solution to the general problem has been given by Krieger [14] for two-sided subshifts. In the one-sided setting, our condition (10) is expressed in terms of the *exit points*  $K_e^c := \{\xi \in K^c : \Theta(\xi) \in K\}$  of  $K$ :

**Theorem 2.** *For a subshift  $K \subset \{0, \dots, J\}^{\mathbb{Z}^-}$ , there is a continuous  $g$ -function with respect to which  $K$  is invariant if and only if*

$$\bigcap_{j=0}^J \overline{(K_e^c)^{*,j}} = \emptyset, \quad (10)$$

where  $^{*,j} : X \rightarrow X$  is given by  $(\dots, x_{-1}, x_0)^{*,j} = (\dots, x_{-1}, x_0 + j \pmod{J})$ .

The reader is assumed to be familiar with MRA wavelet constructions or find the QMF functions for which (a) and (b) hold to be objects worth studying in their own right. The problem of characterizing them presented here is consistent with the formulation in the article [10] that inspired this paper, relying on the characterization of scaling functions by Hernández and Weiss [12] and a procedure for taking square-roots described by The Wutam Consortium [19]. The paper is organized as follows: In the next section, we restate the main question in terms of the Markov process on  $\{0, 1\}^{\mathbb{Z}^-}$  with transition function  $p \circ \tau$ . Our symbolic description of sample paths leads to the definition of an exit point. We prove that the exit points of a subshift define a closed set precisely when the subshift is of finite type. Having seen the role that invariance plays in the symbolic case in Section 3, we return to the unit interval in Section 4, where Theorem 1 is proved. Section 4 also contains a characterization of invariant sets for continuous transition functions on the unit interval. In Section 5 we prove Theorem 2 and extend the result to general subsets (Remark 3). We also show that a subshift has a “strict”  $g$ -function (in the sense of [14]) if and only if it is of finite type. Section 6 concludes.

Two remarks before we proceed. First, for  $L^2(\mathbb{R})$  scaling functions, the QMF condition only requires  $p(\xi/2) + p(\xi/2 + 1/2) = 1$  to hold for a.e.  $\xi$  (see [12, p. 54]). If the Fourier transform of the scaling function is continuous, then so is  $p$ , and we get equality for all  $\xi$ . Since  $\tau : X \rightarrow [0, 1]$  is continuous,  $p \circ \tau$  is continuous

if  $p$  is continuous. By allowing  $p$  to be discontinuous at binary rational  $\xi \in (0, 1)$ , we include such basic examples of QMF functions as the Shannon filter (Example 3), the indicator function of the set  $[0, 1/4] \cup [3/4, 1]$ . Second, (b) is almost trivial. For if only  $p$  is, say, Hölder continuous at 0, so that  $\hat{\Phi}_p$  is continuous at 0, then (b) is simply the hypothesis  $p(0) = 1$ . The subtle point is the fact that the equality in (a) does not have to hold for all  $\xi \in [0, 1]$ . Even in the case when  $p$  is continuous and Hölder continuous at 0, which means that  $\hat{\Phi}_p$  is continuous, we may have  $\sum_{k \in \mathbb{Z}} \hat{\Phi}_p(\xi + k) = 1$  a.e., but  $\sum_{k \in \mathbb{Z}} \hat{\Phi}_p(\xi + k) = 0$  on a closed,  $\theta$ -invariant subset.

## 2. SYMBOLIC DYNAMICS OF SAMPLE PATHS

For every  $\xi_0 \in [0, 1]$  and  $p \in \mathcal{Q}$  there is a stochastic process  $\xi_0, \xi_1, \dots$  with state space  $[0, 1]$  such that

$$\xi_{t+1} = \xi_t/2 \text{ or } \xi_t/2 + 1/2, \quad (11)$$

where  $p(\xi_t/2)$  gives the probability, conditioned on  $\xi_t$ , that  $\xi_{t+1} = \xi_t/2$ . In [9], the *path space* for these processes is the set  $X^+$  of all sequences  $x^+ = (x_1, x_2, \dots)$ , each  $x_i \in \{0, 1\}$ . Sample paths from  $\xi_0 \in [0, 1]$  are recursively defined via

$$\xi_t = \xi_{t-1}/2 + x_t/2, t \geq 1. \quad (12)$$

The binary representation of  $\xi_t(x^+)$  is then given by  $(\xi_0, x_1, \dots, x_t)$ , where  $\xi_0 \in X$  is the binary representation of  $\xi_0 \in [0, 1]$ . Given  $\xi_0 = (\dots, x_{-1}, x_0) \in X$ , let

$$\xi_t(x^+) := (\xi_0, x_1, \dots, x_t), t \geq 1. \quad (13)$$

The set of limit points of (13) is a closed,  $\Theta$ -invariant subset of  $X$  for each  $x^+ \in X^+$ . In symbolic dynamics, such a set is called a subshift or a shift-space. Actually by a subshift one often means the restriction  $\Theta|_K$  of  $\Theta$  to a closed, shift-invariant subset, or the dynamical system  $(K, \Theta|_K)$ . Our purpose here is to study the behavior of  $\xi_t, t \geq 1$ , for fixed  $\xi_0$  rather than the iterates of  $\Theta$ . When sample paths are directed by a probability  $P_g(|\xi_0)$  determined by  $g = p \circ \tau$ , condition (a) becomes equivalent to that  $\xi_t$  converges to  $\mathbf{0} = (\dots, 0, 0)$  or  $\mathbf{1} = (\dots, 1, 1)$  in the product topology,  $P_g(|\xi_0)$ -almost surely for  $d\xi$ -almost every  $\xi_0 \in X$ , where  $d\xi$  is the  $(1/2, 1/2)$ -product measure. If  $B \subset (0, 1)$  is closed,  $\theta$ -invariant and invariant for  $p$ , then  $\tau^{-1}(B)$  is a  $g$ -invariant subshift, necessarily of measure zero by the ergodicity of  $\Theta$  with respect to  $d\xi$ .

To keep the article self-contained, the next subsection introduces notation for subshifts and recalls some of their basic properties. For more information, see for example Lind and Marcus [17].

**2.1. Preliminaries.** Let  $J$  be a positive integer and  $\mathcal{A}$  the alphabet  $\{0, 1, \dots, J\}$ . A *word* over  $\mathcal{A}$  is a finite string  $\bar{x} = (x_i, \dots, x_{i+n})$  of symbols from  $\mathcal{A}$ . We emphasize the length of a word by writing  $\bar{x}_n$  and let  $\bar{j}_n = (j, \dots, j)$  denote the word formed by  $n$  repetitions of the symbol  $j$ . Let

$$X = \mathcal{A}^{\mathbb{Z}-}$$

be the set of all sequences  $\xi = (\dots, x_{-1}, x_0)$  of symbols from  $\mathcal{A}$ . Given  $\xi \in X$  and  $i \leq j \leq 0$ , we define  $\xi^{[i,j]} = (x_i, \dots, x_j)$  and say that  $\bar{x}'$  occurs in  $\xi$  if  $\xi^{[i,j]} = \bar{x}'$  for some  $i \leq j \leq 0$ . The word  $\xi^{[-n+1,0]}$  is called the *initial* word in  $\xi$  of length  $n$ .

*Cylinders* are subsets of  $X$  of the type  $\{\xi \in X : x_{i_k} = x_{i_k}, k = 1, \dots, r\}$ ,  $r \geq 1$ . The cylinder defined by the word  $\bar{x}'_n$  is the set

$$C(\bar{x}'_n) = \{\xi \in X : \xi^{[-n+1, 0]} = \bar{x}'_n\}.$$

Cylinders are defined similarly if we consider  $X^+ = \mathcal{A}^{\mathbb{N}}$  or the two-sided space  $X^\pm = \mathcal{A}^{\mathbb{Z}}$ , points of which are written  $x^+ = (x_1, x_2, \dots)$  and  $x^\pm = (\dots x_{-1}, x_0, x_1, \dots)$ , respectively. In each of the three cases the cylinders give a basis for the product topology ( $\mathcal{A}$  is given the discrete topology), which is compact, that consists of clopen (closed and open) sets. We will consider  $X$  as a compact metric space by defining a metric that is compatible with the topology, given by

$$\rho(\xi, \xi') = \begin{cases} 0 & \text{if } \xi = \xi', \\ 2^{-l(\xi, \xi')} & \text{if } \xi \neq \xi', l(\xi, \xi') = \min\{|j| : x_j \neq x'_j\}. \end{cases} \quad (14)$$

For  $\xi \in X$  and  $E \subset X$ , we let

$$\rho_E(\xi) = \inf_{\xi' \in E} \rho(\xi, \xi').$$

Probability measures on  $\mathcal{A}, X, X^+$  and  $X^\pm$  are assumed to be defined on sigma-fields containing all open subsets. Maps between the spaces will be measurable with respect to the sigma-fields generated by the open sets.

By a subshift of  $X$  we always mean a closed, nonempty proper subset  $K \subset X$  such that  $K = \{\Theta(\xi') : \xi' \in K\}$ , where

$$\Theta(\dots, x_{-1}, x_0) := (\dots, x_{-2}, x_{-1}).$$

Every subshift arises as the set  $X_{\mathcal{F}}$  of all  $\xi \in X$  in which no word of  $\mathcal{F}$  occurs, where  $\mathcal{F}$  is some nonempty collection of “forbidden” words. Conversely, if  $X_{\mathcal{F}}$  is nonempty, it is a subshift. Subshifts of  $X^\pm$  are defined by forbidden words in the same way. We say that  $\mathcal{F}$  *generates*  $X_{\mathcal{F}}$  and that  $\bar{x}$  is *forbidden* from  $X_{\mathcal{F}}$  if  $\bar{x}$  contains a word of  $\mathcal{F}$ . Words that are not forbidden are said to be *allowed*. A subshift is of *finite type* if it is generated by some finite  $\mathcal{F}$ , and *transitive* if for each pair  $\bar{x}', \bar{x}$  of allowed words there is a word  $\bar{x}''$  such that  $(\bar{x}, \bar{x}'', \bar{x}')$  also occurs in a point of the subshift.

The next subsection describes the long-term behavior of sample paths from an initial point  $\xi_0 \in X$ . The description is purely symbolic (probability comes in Section 3) and uses Doeblin’s notion of coupling. The definition of an exit point is very natural in this context, yet the author has been unable to find it elsewhere.

**2.2. Coupling, exit points.** Let  $K \subset X = \{0, 1, \dots, J\}^{\mathbb{Z}-}$  be a subshift and consider a sample path  $\xi_t(x^+) = (\xi_0, x_1, \dots, x_t), t \geq 1$ , from  $\xi_0 \in X$ . If  $\xi_0 \in K^c$ , so that  $\xi_0$  contains a word forbidden from  $K$ , then  $\xi_t$  contains a word forbidden from  $K$  for all  $t \geq 1$ , where  $\rho(\xi_t, K) \geq 2^{-t} \rho(\xi_0, K)$ . Let us say that  $\xi_0, \xi_1, \dots$

couples with  $K$  to  $n$  if  $\rho(\xi_t, K) = 2^{-t} \rho(\xi_0, K)$  for  $t = 1, \dots, n$ ,

couples with  $K$  if  $\rho(\xi_t, K) = 2^{-t} \rho(\xi_0, K)$  for all  $t \geq 1$ ,

decouples from  $K$  at time  $t$  if  $\rho(\xi_t, K) > \rho(\xi_{t-1}, K)$ ,

converges to  $K$ , written  $\xi_t \rightarrow K$ , if  $\limsup_{t \rightarrow \infty} \rho(\xi_t, K) = 0$ .

If  $\xi_0 \in K$ , then the sample path either stays in  $K$ , or it permanently leaves  $K$  at some time  $t \geq 1$ . The exit point  $\xi_t \in K^c$  satisfies  $\Theta(\xi_t) \in K$ .

**Definition 1.** Let  $E$  be a subset of  $X = \{0, 1, \dots, J\}^{\mathbb{Z}^-}$ . A point  $\xi \in X$  is a point of exit from  $E$  if  $\xi \in E^c$  and  $\Theta(\xi) \in E$ .

We will call a point of exit from  $E$  simply an *exit point* of  $E$ . To emphasize that exit points of a set are points in the complement of that set, the set of exit points of  $E \subset X$  is denoted by  $E_e^c$ . We can write  $E_e^c = \Theta^{-1}(E) \cap E^c$ . Note that  $E_e^c \neq \emptyset$  if  $E$  is a closed proper subset of  $X$ . (Indeed, if  $E \subset X$  has no exit points, then we must have  $(\xi, \bar{x}') \in E$  for every  $\xi \in E$  and every word  $\bar{x}'$  over  $\{0, \dots, J\}$ , which means that  $E$  is dense. If  $E$  is closed, this implies that  $E = X$ .) In fact for each  $\xi \in E$  there is then a word  $\bar{x}'$  such that  $(\xi, \bar{x}') \in E_e^c$ .

We prove below that for a subshift  $K \subset X$ ,  $K_e^c$  is closed if and only if  $K$  is of finite type. For this it will be convenient to have versions of the maps  $^{*,j}$  in (10) defined on words:

$$(x_1, \dots, x_n)^{*,j} := (x_1, \dots, x_n + j \pmod{J}), j \in \mathbb{Z}.$$

When  $J = 1$ , the map  $\xi \mapsto \xi^* := \xi^{*,1}$  models addition by  $1/2 \pmod{1}$  on the unit interval  $[0, 1]$ . In this case  $K_b := (K_e^c)^* := \{\xi \in X : \xi^* \in K_e^c\}$  coincides with the set of  $\xi \in K$  that satisfy  $\xi^* \in K^c$  (Proposition 1 (a)). Call them *barrier points* (cf. [9, p. 76]). For  $J \geq 1$ , we define  $K_b = \{\xi \in K : \xi^{*,j} \in K^c, 1 \leq j \leq J\}$ . Thus  $(\xi, j') \in K_b$  if  $\xi \in K$  and  $(\xi, j)$  is a point of exit from  $K$  for all  $j \in \{0, 1, \dots, J\} \setminus \{j'\}$ .

**Remark 1.** (i) In [9], a barrier point is a point  $\xi \in K$  with  $\xi^* \in K^c$  and  $g(\xi) = 1$ . Here barrier points are defined without reference to a  $g$ -function. (ii) While  $K_b = (K_e^c)^*$  if  $J = 1$ ,  $K_b$  may be empty if  $J \geq 2$ .

**Proposition 1.** Let  $K \subset \{0, 1, \dots, J\}^{\mathbb{Z}^-}$  be a subshift.

- (a) When  $J = 1$ ,  $K_b = \{\xi \in K : \xi^* \in K^c\}$  and  $K_e^c = \{\xi \in K^c : \xi^* \in K\}$
- (b)  $K_b$  is closed if  $K$  is of finite type.
- (c) The following are equivalent:
  - (i)  $K$  is of finite type.
  - (ii)  $K_e^c$  is closed.
  - (iii)  $K \cap \overline{K_e^c} = \emptyset$ .

*Proof.* Write  $K = X_{\mathcal{F}}$  for some collection of words  $\mathcal{F}$ . Assume without loss of generality that no word in  $\mathcal{F}$  contains another word from  $\mathcal{F}$  as a subword; i.e., if  $\bar{x} = (x_{-n}, \dots, x_0) \in \mathcal{F}$ , then  $(x_{-n+1}, \dots, x_0) \notin \mathcal{F}$  and  $(x_{-n}, \dots, x_{-1}) \notin \mathcal{F}$ .

(a) Suppose that  $\xi \in K$  and  $\xi^* \in K^c$ . Then  $\xi^*$  satisfies  $(\xi^*)^* = \xi \in K$  and  $\Theta(\xi^*) = \Theta(\xi) \in K$  by  $\Theta$ -invariance of  $K$ . (We have  $\Theta(\xi'^*) = \Theta(\xi')$  for every  $\xi' \in X$ .) Thus  $\xi^* \in K_e^c$ , i.e.,  $\xi \in (K_e^c)^*$ . Since  $\xi \in \{\xi \in K : \xi^* \in K^c\}$  was arbitrary,  $\{\xi \in K : \xi^* \in K^c\} \subset (K_e^c)^*$ . For the reverse inclusion, let  $\xi \in (K_e^c)^*$ , so that  $\xi^* \in K^c$  and  $\Theta(\xi^*) \in K$ . Then  $\xi \in K$ , for otherwise  $\Theta(\xi^*)$  would not be on the form  $\Theta(\xi')$  for some  $\xi' \in K$ . Since  $\xi \in (K_e^c)^*$  was arbitrary,  $(K_e^c)^* \subset \{\xi \in K : \xi^* \in K^c\}$ , so  $K_b = \{\xi \in K : \xi^* \in K^c\}$ . Finally, that  $K_e^c = \{\xi \in K^c : \xi^* \in K\}$  follows from that  $K_b = \{\xi \in K : \xi^* \in K^c\}$ .

(b) Let  $\mathcal{B}$  be the collection of words  $\bar{x} \notin \mathcal{F}$  such that  $\bar{x}^{*,j} \in \mathcal{F}$  for all  $j \in \{1, \dots, J\}$ . We have  $K_b = K \cap (\cup_{\bar{x} \in \mathcal{B}} C(\bar{x}_n))$ . If  $\mathcal{F}$  is finite, then so is  $\mathcal{B}$ , which means that  $K \cap (\cup_{\bar{x} \in \mathcal{B}} C(\bar{x}_n))$  is closed.

(c) (i) implies (ii): We can write

$$K_e^c = \left( \bigcup_{i=1}^J K^{*,i} \right) \cap \left( \bigcup_{\bar{x} \in \mathcal{F}} C(\bar{x}) \right). \quad (15)$$

To see this, we first note that every point of  $K_e^c$  can be written  $\xi^{*,j}$  for some  $\xi \in K$  and  $j \in \{1, \dots, J\}$ . (This follows from that  $K = \{\Theta(\xi') : \xi' \in K\}$  and  $\Theta(\xi) \in K$  if  $\xi \in K_e^c$ . Indeed, if  $\xi \in K_e^c$  and we had  $\xi^{*,j} \in K^c$  for  $j = 1, \dots, J$ , then  $\Theta(\xi)$  would not be on the form  $\xi = \Theta(\xi')$  for some  $\xi' \in K$ .) So  $K_e^c \subset \cup_{i=1}^J K^{*,i}$ . Given  $\xi \in \cup_{i=1}^J K^{*,i}$ , we have  $\xi \in K_e^c$  if and only if  $\xi^{[-n+1,0]} \in \mathcal{F}$  for some  $n$ . Thus  $K_e^c = (\cup_{j=1}^J K^{*,j}) \cap (\cup_{\bar{x} \in \mathcal{F}} C(\bar{x}))$ , i.e. (15) holds. Now, for  $j \in \{0, 1, \dots, J\}$ ,  $K^{*,j}$  is closed by the continuity of  $^{*,j}$ . Since  $C(\bar{x})$  is closed for every  $\bar{x} \in \mathcal{F}$ ,  $\cup_{\bar{x} \in \mathcal{F}} C(\bar{x})$  is closed if  $\mathcal{F}$  is finite. From this and (15) it follows that  $K_e^c$  is closed if  $\mathcal{F}$  is finite. (ii) immediately implies (iii) since  $K_e^c \subset K^c$  by definition. (iii) implies (i): Assume that  $K \cap \overline{K_e^c} = \emptyset$  and suppose for contradiction that  $\mathcal{F}$  is infinite. For each  $i \geq 1$  we can then choose a word  $\bar{x}_{n_i} = (x_{-n_i+1}, \dots, x_0) \in \mathcal{F}$  of length  $n_i \geq i$ . By our assumption on  $\mathcal{F}$ , both  $(x_{-n_i+2}, \dots, x_0)$  and  $(x_{-n_i+1}, \dots, x_{-1})$  are allowed. Thus we can take  $\xi(i) \in K$  such that  $(x_{-n_i+1}, \dots, x_{-1})$  is the initial word in  $\xi(i)$  of length  $n_i - 1$ . Then the initial word in  $\xi'(i) := (\xi(i), x_0)$  of length  $n_i - 1$  is allowed, but the initial word in  $\xi'(i)$  of length  $n_i$  is forbidden. That is,  $\xi'(i) \in K_e^c$  and  $\rho(\xi'(i), K) \leq 2^{-n_i+1}$ . Because  $X$  is a compact metric space,  $\xi'(i), i \geq 1$ , has a convergent subsequence  $\xi'(i_k), k \geq 1$ . For this subsequence, we have  $\rho(\xi'(i_k), K) \rightarrow 0$ , hence  $\lim_{k \rightarrow \infty} \xi'(i_k) \in K \cap \overline{K_e^c}$ .  $\square$

### 3. INVARIANCE

A probability measure on the set of sequences  $x^\pm = (\dots, x_{-1}, x_0, x_1, \dots)$  with entries in  $\{0, 1\}$  gives rise to two well-studied random processes. The  $X$ -valued process  $\xi_t = (\dots, x_{t-1}, x_t), t \in \mathbb{Z}$ , has the Markov property given any probability distribution for  $x^\pm \in X^\pm$ . This is immediate from that for each  $t \in \mathbb{Z}$ , the sigma-field  $\sigma(\xi_t, \xi_{t-1}, \dots) = \sigma(\xi_t)$ . In contrast, the coordinate process  $x_t, t \in \mathbb{Z}$ , which takes values in  $\{0, 1\}$ , typically has “infinite connections”. That is, the transition probabilities depend on the entire past history of the process.

The Markov process is determined by its transition function and initial distribution. In particular, for  $p \in \mathcal{Q}$  and  $\xi_0 \in X$ , there is a unique probability  $P_g(\cdot | \xi_0)$  on  $\{0, 1\}^\mathbb{N}$  such that  $\xi_t, t \geq 1$ , depends on the past according to  $g := p \circ \tau$ . (Conditioned on  $\xi_{t-1}$ ,  $\xi_t$  takes the value  $(\xi_{t-1}, j)$  with probability  $g(\xi, j), j = 0, 1$ .) As mentioned above, the probabilistic interpretation of the condition (a) in (5) says that

$$(a) \ P_g(\xi_t \rightarrow \{0, 1\} | \xi_0) = 1 \text{ for } d\xi\text{-almost every } \xi_0 \in X.$$

From [9, Theorem 11.1] we know that (a) depends on the attracting properties of any invariant sets besides  $\{0\}$  and  $\{1\}$ . (These sets are always invariant by virtue of the fact that  $p(0) = p(1) = 1$  implies  $g(0) = g(1) = 1$ .) In general, if  $K$  is  $g$ -invariant, then the probability of coupling with  $K$  to  $n$  tends to 1 as  $\rho(\xi_0, K) \rightarrow 0$ . If  $g$  has “summable variations”, meaning that the sequence

$$\text{var}_n(g) := \sup\{|g(\xi) - g(\xi')| : \rho(\xi, \xi') < 2^{-n}\}, n \geq 1,$$

satisfies

$$\sum_{n=1}^{\infty} \text{var}_n(g) < +\infty,$$

then

$$P_g(\xi_t \text{ couples with } K | \xi_0) \rightarrow 1 \text{ as } \rho(\xi_0, K) \rightarrow 0. \quad (16)$$



Say that the invariant set  $K$  is a *strong attractor* if (16) is satisfied, and that  $K$  is *inaccessible* if

$$P_g(\xi_t \rightarrow K | \xi_0) = 0 \text{ for every } \xi_0 \in K^c. \quad (17)$$

While (a) is equivalent to (a), (b) is implied by

$$(b) \{0\} \text{ and } \{1\} \text{ are strong attractors.} \quad (18)$$

If (b) is satisfied, then (a) holds if there are no invariant sets disjoint from  $\{0\}$  and  $\{1\}$ . Conversely, if  $g$  has summable variations and  $K \subset X \setminus \{0, 1\}$  is  $g$ -invariant, then (16) has the effect that the equality in (a) fails on a set of positive measure. But (a) does not rule out the presence of additional invariant sets if the  $g$ -function is only supposed to be continuous.

A caricatural example is discussed in [10, p. 1102-1103] starting from  $p(\xi) = \cos^2(3\pi\xi)$ , a QMF function with  $p(1/3) = p(2/3) = 1$ . With respect to such  $p$ , the  $\theta$ -invariant set  $B = \{1/3, 2/3\}$  is invariant, where

$$K := \tau^{-1}(B) = \{(\dots, 0, 1, 0) \text{ and } (\dots, 1, 0, 1)\} = X_{\{\bar{0}_2, \bar{1}_2\}}$$

is invariant for  $p \circ \tau$ . To prevent (16),  $p$  is given sharp cusps at  $1/3$  and  $2/3$ , with corresponding modifications near  $1/6$  and  $5/6$  to retain the QMF condition. Sample paths are still attracted to  $K$ , but they never “accelerate fast enough to converge”, to quote Gundy’s description of the effect of an inaccessible invariant set. The typical sample path couples with  $K$  to some finite  $n$ , decouples, possibly recouples a finite number of times, then succumbs to the strong attraction of  $\{0\}$  and  $\{1\}$ . How generic is this behavior? Given a closed,  $\Theta$ -invariant set  $K$  that does not contain  $0$  or  $1$ , is there a continuous  $g \in \mathcal{G}$  for which (a) and (b) hold, where  $K$  is invariant but inaccessible?

Before giving the answer to this question we discuss the examples  $K = X_{\{\bar{0}_n, \bar{1}_n\}}$  from [9, Section 13]. For fixed  $n \geq 3$ ,  $K = X_{\{\bar{0}_n, \bar{1}_n\}}$  is a closed, totally disconnected perfect subset of  $X$ , so it is homeomorphic to the Cantor set. Suppose that we have defined  $g \in \mathcal{G}$  so that  $g$  is continuous, (b) holds and  $K$  is  $g$ -invariant, but that  $g$  has no zeros outside  $K_e^c$  besides the zeros at  $0^*$  and  $1^*$ , which are required for  $g(0) = g(1) = 1$ . (Such a construction is always possible when  $K$  is of finite type; see Theorem 3.) To prevent sample paths from initial points in the complement of  $K$  to converge to  $K$ , we increase the rate at which they decouple. That  $\xi_t, t \geq 0$ , decouples from  $K$  at time  $t$  means that  $\xi_t \in U_e$ , where

$$U_e := C(\bar{0}_n) \cup C(\bar{1}_n).$$

By Levy’s conditional Borel-Cantelli Lemma (see for example Chen [2] or [9, Lemma 4.1’]), we will have  $\xi_t \in U_e$  for infinitely many values of  $t \geq 1$  if

$$\sum_{t=0}^{\infty} P(\xi_{t+1} \in U_e | \xi_t) = +\infty, P(|\xi_0)\text{-a.s.} \quad (19)$$

The initial words  $\bar{0}_{n-1}$  and  $\bar{1}_{n-1}$  are “crucial” in that  $U_e$  can be reached in one step. By our assumptions on  $g$ , the probability to reach  $U_e := C(\bar{0}_{n-1}) \cup C(\bar{1}_{n-1})$  in no more than  $n-2$  steps is positive for every  $\xi_0 \in X$ . (If  $\xi_0 = 0$  or  $1$ , no steps have to be taken. If  $\xi_0 \in K$ , there is a path from  $\xi_0$  to  $U_e$  lying entirely in  $K$ , where  $g$  is strictly positive. Finally, if  $\xi_0 \in (K \cup \{0, 1\})^c$ , then  $(\xi_0, \bar{x}) \notin K_e^c \cup \{0^*, 1^*\}$  for any  $\bar{x}$ , so that every finite-step transition probability is positive.) The strictly positive finite-step transition probability is a continuous function of  $\xi_0$ , so it is



bounded away from zero. This means (essentially by the renewal theorem) that we can find  $\beta > 0$ , not depending on  $\xi_0 \in X$ , such that the recurrence times  $t_1, t_2, \dots$  for crucial words (i.e., the times when  $\xi_t \in U_c$ ) satisfy  $t_j \leq \beta j$ ,  $P(|\xi_0)$ -a.s. Setting  $g = |\log_2 \rho(\xi, K_e^c)|^{-1}$  on  $U_e \setminus K_e^c$ , with a corresponding modification on  $U_e^* \setminus K_b$ , we get

$$P(\xi_{t_j+1} \in U_e | \xi_{t_j}) \geq \frac{1}{l + t_j} \geq \frac{1}{l + \beta j},$$

where  $l$  is the integer with  $\rho(\xi_0, K) = 2^{-l}$ . (Here we have used that  $\rho(\xi_0, K) = 2^{-l}$  implies  $\rho(\xi_t, K_e^c) \geq 2^{-(t+l)}$ , which follows once we observe that the initial word of length  $t+l$  in  $\xi_t$  cannot be the initial word of an exit point of  $K$  if the initial word in  $\xi_0$  of length  $l$  is forbidden from  $K$ .) Because (19) holds,  $K$  is inaccessible. Using that  $\{0\}$  and  $\{1\}$  are strong attractors and that there are no other invariant sets, we can show that the equality in (a) holds on  $K^c$ , hence almost everywhere.<sup>2</sup>

The above construction and argument relies (only) on the assumption that  $K$  is of finite type. If  $K$  is  $g$ -invariant but not of finite type, then  $g$  must vanish not only on  $K_e^c$ , but on  $\overline{K_e^c} \cap K$  as well (Proposition 1(c)). This may leave us without a lower bound on the probability to encounter a crucial word in any number of steps. For the subshift in Example 1, sample paths from  $\xi_0 = (\dots, 0, 1, 0)$  or  $(\dots, 1, 0, 1)$  will never encounter a crucial word. However, as long as (b) holds and the zeros of  $g$  are contained in  $\overline{K_e^c} \cup C((0, 1)) \cup C((1, 0))$ ,  $\{0, 1\}$  is still accessible from any  $\xi_0 \in K^c$  in the sense that  $P(\rho(\xi_k, \{0, 1\}) \leq 2^{-k} | \xi_0) > 0$  for each  $k \geq 1$ . Consider therefore a sequence of trials, where trial  $n \geq 0$  consists of the attempt to reach  $U_{0,1} = C(\bar{0}_{k+1}) \cup C(\bar{1}_{k+1})$  by  $k$  consecutive steps towards either  $0$  or  $1$ , depending on whether the initial symbol in  $\xi_{nk}$  is  $0$  or  $1$ . For  $k$  so large that  $U_{0,1}$  is disjoint from  $K$  and  $g(\xi) = |\log_2 \rho(\xi, \overline{K_e^c})|^{-1/k}$  on a neighborhood  $U_e$  of  $\overline{K_e^c}$ , we obtain (below), for some  $\lambda' > 0$  and all  $n \geq 1$ , that the success probability

$$P_g(\xi_{nk+k} \in U_{0,1} | \xi_{nk}) \geq \frac{\lambda'}{l + nk + k}, \text{ where } l = |\log_2 \rho(\xi_0, K)|. \quad (20)$$

Setting  $g \equiv 1$  on  $U_{0,1}$  achieves (a) since  $U_{0,1}$  is visited infinitely often if  $\xi_0 \in K^c$ , again by Borel-Cantelli.

A construction of the second type is always possible if  $K \subset X \setminus \{0, 1\}$  satisfies

$$\overline{K_e^c} \cap \overline{K_b} = \emptyset. \quad (21)$$

This is also a necessary constraint if we require that  $g$  be continuous, for then  $g$ -invariance means  $g(\xi) = 0$  for all  $\xi \in \overline{K_e^c}$  and hence  $g(\xi) = 1$  for all  $\xi \in \overline{K_b}$ . The construction does not prove Theorem 1 however, because it need not be the case that  $g$  is on the form  $g = p \circ \tau$  for some continuous  $p \in \mathcal{Q}$ . Before we go back to the unit interval and the proof of this theorem, we give one example to show that condition (21) does not require that  $K$  be finitely generated. (The condition is satisfied for subshifts of finite type by Proposition 1(c).)

**Example 1.** Let  $(0, 1, 0)^1 = (0, 1, 0)$  and define words of length  $3n$ ,  $n \geq 2$ , recursively via  $(0, 1, 0)^n = ((0, 1, 0)^{n-1}, 0, 1, 0)$ . The words  $(1, 0, 1)^n$ ,  $n \geq 1$ , are defined

<sup>2</sup>We remark that in [9, Lemma 13.2] it is falsely claimed that the recurrence times for barrier words form an iid sequence. The proof of Theorem 13.1 can be corrected by introducing the notion of a crucial word and arguing as above.

similarly. If

$$\mathcal{F} = \{(0, 0, 0, 0), (1, 1, 1, 1)\} \bigcup_{n=1}^{\infty} \{(0, (0, 1, 0)^n, 0), (1, (1, 0, 1)^n, 1)\}, \quad (22)$$

then  $K = X_{\mathcal{F}}$  is not of finite type since  $(\dots, 0, 1, 0, 1, 1)$  and  $(\dots, 1, 0, 1, 0, 0)$  are points of  $K \cap \overline{K_e^c}$ . That (21) holds follows from that  $\overline{K_e^c} \subset C((0, 0)) \cup C((1, 1))$  and  $(\overline{K_e^c})^* \subset C((0, 1)) \cup C((1, 0))$ .

#### 4. TRANSITION FUNCTIONS ON THE UNIT INTERVAL

In this section we study subsets of the unit interval that are invariant<sup>3</sup> for the process that we have considered. Examples have been provided by, among others, Conze and Raugi [5, Section VII], Cohen and Conze [4] and Gundy [9, Section 13]. Section 4.3 provides necessary and sufficient conditions for a set  $B \subset [0, 1]$  to be invariant with respect to a continuous transition function. Section 4.4 contains the proof of Theorem 1. The next subsection clarifies notation from the introduction.

**4.1. Some definitions.** The definitions in this subsection are standard with the possible exception for the map in (24).

**4.1.1. Binary representations.** The map

$$(\dots, x_{-1}, x_0) \mapsto \tau(\dots, x_{-1}, x_0) := \sum_{j=-\infty}^0 x_j 2^{j-1}$$

is continuous and takes  $X$  onto  $[0, 1]$ . It becomes a measure-preserving isomorphism from  $X$  to the Borel-Lebesgue unit interval when  $X$  is supplied with the infinite product of the measure  $(1/2, 1/2)$ . If  $\tau(\xi) = \xi$ , then  $\xi \in X$  is called a binary representation of  $\xi$ . If  $\xi \in (0, 1)$  is binary rational (that is, if  $\xi = k/2^n$  for some integers  $k > 0$  and  $n$  with  $k/2^n < 1$ ), then  $\tau^{-1}(\xi)$  consists of two sequences: one having a finite number of zeros, the other having a finite number of ones. If  $\xi \in (0, 1)$  is not binary rational, it has a unique binary representation.

**4.1.2. Multiplication by 2 and addition by  $1/2 \pmod{1}$ .** As a map of  $[0, 1)$  to itself, the doubling map (8) is defined

$$\theta(\xi) := \begin{cases} 2\xi & \text{if } \xi \in [0, 1/2), \\ 2\xi - 1 & \text{if } \xi \in [1/2, 1). \end{cases} \quad (23)$$

The map  $\xi \mapsto \xi^* = \xi + 1/2 \pmod{1}$  is unambiguously defined on  $[0, 1)$ . We define  $\xi^*$  for all  $\xi \in [0, 1]$  as follows:

$$\xi^* := \begin{cases} \xi + 1/2 & \text{if } \xi \in [0, 1/2], \\ \xi - 1/2 & \text{if } \xi \in (1/2, 1]. \end{cases} \quad (24)$$

The corresponding map on  $X = \{0, 1\}^{\mathbb{Z}-}$  has the property that  $(\xi^*)^* = \xi$  for all  $\xi \in X$ . We do have  $(\xi^*)^* = \xi$  if  $\xi \in (0, 1]$ , but  $(0^*)^* = (1/2)^* = 1$ . In analogy with the symbolic map,

$$(\overline{E})^* = \overline{E^*} \quad (25)$$

for any  $E \subset [0, 1]$ . Here  $E^* := \{\xi^* : \xi \in E\}$ .

---

<sup>3</sup>Some authors use the term “stochastically closed”.

4.1.3. *The distance function.* For  $\xi \in [0, 1]$  and an arbitrary set  $E \subset [0, 1]$ , we let

$$d_E(\xi) := \inf_{\xi' \in E} |\xi - \xi'|. \quad (26)$$

4.2. **Transition functions.** In this subsection and the next we drop the requirement that  $p(0) = 1$  and call  $p: [0, 1] \rightarrow [0, 1]$  a transition function if

$$p(\xi/2) + p(\xi/2 + 1/2) = 1 \text{ for every } \xi \in [0, 1]. \quad (27)$$

Note that (27) implies  $p(0) = 1 - p(1/2) = p(1)$ . Replacing  $\xi$  in (27) by  $2\xi$ , (27) can be rewritten as

$$p(\xi) + p(\xi + 1/2) = 1 \text{ for every } \xi \in [0, 1/2]. \quad (28)$$

Alternatively,

$$p(\xi) + p(\xi^*) = 1 \text{ for every } \xi \in \mathcal{X}, \quad (29)$$

where  $\mathcal{X}$  is any subset of  $[0, 1]$  with  $\mathcal{X}^* \cup \mathcal{X} = [0, 1]$ . In particular, the above conditions are all equivalent to that

$$p(\xi) + p(\xi^*) = 1 \text{ for every } \xi \in [0, 1]. \quad (30)$$

4.3. **Invariant sets.** We are here interested in conditions that ensure the existence of a transition with respect to which a given set  $B \subset [0, 1]$  is invariant, that is, for which  $P_p(\xi_1 \in B | \xi_0) = 1$  for all  $\xi_0 \in B$ . The latter condition holds if and only if  $p(\xi) = 0$  for all  $\xi \in B_e^c$ , where

$$B_e^c := \{\xi/2 + j/2 : \xi \in B, j \in \{0, 1\}\} \cap B^c. \quad (31)$$

If  $B \subset [0, 1]$ , then  $B_e^c = \theta^{-1}(B) \cap B^c$ . In order for such a set to be invariant with respect to some  $p$ , we must have  $B \subset \theta(B)$ . Indeed,  $B \subset \theta(B)$  means that at least one of  $\xi/2$  and  $\xi/2 + 1/2$  is in  $B$  if  $\xi \in B$ , which is clearly a necessary condition for processes that start in  $B$  to be able to remain in  $B$ . We have  $B \subset \theta(B)$  if and only if  $B_e^c \cap B_b = \emptyset$  (Lemma 1(a)), where

$$B_b := (B_e^c)^*. \quad (32)$$

If  $p(\xi) = 0$  for all  $\xi \in B_e^c$ , then  $p(\xi) = 1$  for all  $\xi \in B_b$ . So for  $B \subset [0, 1]$  to be invariant with respect to a continuous  $p$ , we must have

$$\overline{B_e^c} \cap \overline{B_b} = \emptyset. \quad (33)$$

**Proposition 2.** *Let  $B \subset [0, 1]$ . There is a continuous transition function with respect to which  $B$  is invariant if and only if (33) is satisfied.*

*Proof.* 'Only if' was discussed above. If (33) holds there is a closed set  $N_e \subset [0, 1]$  with  $\overline{B_e^c} \subset N_e \subset [0, 1] \setminus \overline{B_b}$ . We want  $N_e$  such that  $N_e \cap N_e^* = \emptyset$ . To achieve this, take  $\delta > 0$  so that  $|\xi - \xi'| > \delta$  if  $\xi \in \overline{B_e^c}$  and  $\xi' \in \overline{B_b}$ . For  $\xi \in \overline{B_e^c}$ , let  $[\xi - \varepsilon, \xi + \varepsilon]$  be a subinterval of  $[0, 1]$  whose length does not exceed  $\delta/2$  and that does not have  $1/2$  as an end-point. Since  $\overline{B_e^c}$  is compact, it is contained in a finite union of these intervals. Let  $N_e$  be such a union. Then  $N_b := N_e^*$  is a finite union of closed intervals whose lengths do not exceed  $\delta/2$ , each containing a point  $\xi \in \overline{B_b}$ . (If  $I = [a, b] \subset [0, 1]$  and  $1/2 \notin \{a, b\}$ ,  $I^*$  is a closed interval unless  $1/2 \in I$ . In this case  $I^*$  is the union of two closed intervals.) Our choice of  $\delta$  gives  $N_e \cap N_b = \emptyset$ .

Now define

$$p(\xi) = \begin{cases} d_{B_e^c}(\xi) & \text{if } \xi \in N_e, \\ 1 - d_{B_e^c}(\xi^*) & \text{if } \xi \in N_b. \end{cases}$$

Then the equality in (30) holds on  $N_e \cup N_b$ . Extend  $p$  continuously to  $[0, 1/2]$  in such a way that  $p(0) = 1 - p(1/2)$ . Setting  $p(\xi) = 1 - p(\xi^*)$  for  $\xi \in (1/2, 1]$ ,  $p$  is continuous on  $[0, 1]$  and satisfies (30). Since  $p \equiv 0$  on  $B_e^c$ ,  $B$  is  $p$ -invariant.  $\square$

A sufficient condition for (33) to hold is of course that  $B_e^c = \emptyset$ . (This is the case, for instance, if  $B$  is the set of binary irrational  $\xi \in (0, 1)$ .) Then  $B$  is invariant for every  $p$ . Below are two examples where  $B_e^c \neq \emptyset$ .

**Example 2.** For the  $\theta$ -invariant set  $B = \{1/3, 2/3\}$ , we have  $B_e^c = \{1/6, 5/6\}$  and  $B_b = \{1/3, 2/3\} = B$ . We have mentioned that  $B$  is invariant for  $p(\xi) = \cos^2(3\pi\xi)$ .

**Example 3.** Let  $B = [0, 1/4] \cup [3/4, 1]$ . Here  $B_e^c = [3/8, 5/8)$  and  $B_b = [1/8) \cup [7/8)$ . The set  $B$  is invariant for, for example, the Shannon-filter (cf. [18, p. 495]), the indicator function of  $B$ .

Lemma 1(c) lists properties of closed subsets of  $(0, 1)$  that are  $\theta$ -invariant. If  $B \subset (0, 1)$  is closed and  $\theta$ -invariant, then  $\overline{B_b} \cup \overline{B_e^c}$  has no binary rationals. This means that for such a  $B$ , (33) remains necessary in the more general case when  $p$  is only required to be continuous at binary irrationals. (This class contains all  $p$  for which  $p \circ \tau$  is continuous on  $\{0, 1\}^{\mathbb{Z}-}$ .) It also means that  $\tau^{-1}(B_e^c) = K_e^c$ , where  $K := \tau^{-1}(B)$ . If  $K$  is of finite type it follows from Proposition 1(c) and the continuity of  $\tau$  that  $B_e^c$  (and hence  $B_b$ ) is closed, so that (33) holds. By Proposition 2, these facts imply that every  $B \subset (0, 1)$  obtained by proscribing a finite number of words from the binary representation of  $\xi \in (0, 1)$  arises as an invariant set for some continuous  $p$ .

**Lemma 1.** *Let  $B \subset (0, 1)$ .*

- (a)  $B \subset \theta(B)$  if and only if  $B_e^c \cap B_b = \emptyset$ .
- (b) If  $B$  is closed, then  $B_e^c$  is nonempty.
- (c) Suppose that  $B$  is closed and  $\theta$ -invariant (i.e.,  $B = \theta(B)$ ). Then
  - (i)  $B$  has Lebesgue measure zero,
  - (ii)  $B_b = \{\xi \in B : \xi^* \in B^c\}$  and  $B_e^c = \{\xi \in B^c : \xi^* \in B\}$ .
  - (iii) None of  $B, \overline{B_e^c}, \overline{B_b}$  contain binary rationals.

*Proof.* (a)  $B \subset \theta(B)$  means that for  $\xi \in B$ , either  $\xi/2$  or  $\xi/2 + 1/2 = (\xi/2)^*$  (or both) is in  $B$ . So if  $B \subset \theta(B)$  fails there exists  $\xi_0 \in B$  such that  $\xi_0/2$  and  $(\xi_0/2)^*$  are both in  $B^c$ . Then  $\xi_0/2$  and  $(\xi_0/2)^*$  are in fact both in  $B_e^c \cap (B_e^c)^*$ . Thus  $B_e^c \cap B_b = \emptyset \Rightarrow B \subset \theta(B)$ . If  $\xi \in B_e^c \cap (B_e^c)^*$ , then  $\xi_0 := \theta(\xi) \in B$ , where  $\xi_0/2$  and  $(\xi_0/2)^*$  are both in  $B^c$ . This means that  $B \subset \theta(B)$  fails. So  $B \subset \theta(B) \Rightarrow B_e^c \cap B_b = \emptyset$ .

(b) For any  $\xi_0 \in B$  and  $x^+ \in \{0, 1\}^{\mathbb{N}}$ , the recursion (12) that defines  $\xi_t(x^+)$  gives

$$\xi_t(x^+) = \frac{\xi_0 + \sum_{i=1}^t x_i 2^{i-1}}{2^t}. \quad (34)$$

If there are no points of exit from  $B$ , then  $\xi_t(x^+) \in B$  for every  $x^+ \in \{0, 1\}^{\mathbb{N}}$  and all  $t \geq 1$ . From (34) we see that  $\{\xi_t(x^+) : x^+ \in \{0, 1\}^{\mathbb{N}}, t \geq 1\}$  is dense in  $[0, 1]$ , so  $B$  is dense in  $[0, 1]$ . Since  $B$  is closed, this implies  $B = [0, 1]$ , a contradiction.

- (b) (i) That the Lebesgue measure of  $B$  is either zero or one follows from that  $\theta$  is ergodic with respect to this measure (see for example Einsiedler and Ward [8, p. 27]). A closed subset of  $(0, 1)$  cannot have Lebesgue measure one.  
(ii) This follows by arguing as in Proposition 1(a) with  $\theta$  in the role of  $\Theta$ .  
(iii) If  $\xi = k/2^n \in (0, 1)$ , then  $\theta^n(\xi) = 0$ . Since  $B = \theta(B)$  and  $B \subset (0, 1)$ ,  $B$  has no binary rationals. The same must be true of  $\overline{B_b} \subset B$  and  $\overline{B_e^c} = (\overline{B_b})^*$ .  $\square$

**4.4. Theorem 1.** The proof of Theorem 1 is divided into two parts. Given a closed  $\theta$ -invariant  $B \subset (0, 1)$  with  $\overline{B_e^c} \cap \overline{B_b} = \emptyset$ , we first construct a continuous  $p \in \mathcal{Q}$  with respect to which  $B$  is invariant. We then verify that  $p$  satisfies (5) and (6). The verification of (6) uses the estimate (36) on the speed at which sample paths can approach  $B_e^c$ .

**Lemma 2.** Let  $B \subset (0, 1)$  be closed and  $\theta$ -invariant and let  $\xi_0 \in B^c$ . There is a constant  $\alpha = \alpha(\xi_0) > 0$  such that for any sample path  $\xi_t = \xi_t(x^+)$ ,  $t \geq 0$ , from  $\xi_0$ ,

$$d_B(\xi_t) \geq \alpha 2^{-t} \text{ for all } t \geq 0, \quad (35)$$

$$d_{B_e^c}(\xi_t) \geq \alpha 2^{-t} \text{ for all } t \geq 1. \quad (36)$$

**Remark 2.** The corresponding estimates on the speed at which sample paths from  $\xi_0 \in \{0, 1\}^{\mathbb{Z}^-}$  can approach a subshift or its exit points do not imply (35) and (36). For there is no  $M > 0$  such that  $|\tau(\xi) - \tau(\xi')| \geq M\rho(\xi, \xi')$  for all  $\xi, \xi' \in X$ .

*Proof.* Pick  $\delta \in (0, 1)$  so that  $|\xi - 1/2| > \delta$  for all  $\xi \in B \cup \overline{B_e^c}$ . (Lemma 1(iii) gives  $1/2 \notin B \cup \overline{B_e^c}$ .) Let  $\xi_0 \in B^c$  be given, let  $x^+ \in \{0, 1\}^{\mathbb{N}}$  and consider the sample path  $\xi_t = \xi_t(x^+)$  defined by the recursion (12). We first show that (35) holds with

$$\alpha := \min(\delta, d_B(\xi_0)).$$

This choice of  $\alpha > 0$  gives  $d_B(\xi_t) \geq \alpha 2^{-t}$  if  $t = 0$ . So it is enough to show that  $d_B(\xi_{t-1}) \geq \alpha 2^{-(t-1)}$  implies  $d_B(\xi_t) \geq \alpha 2^{-t}$  for all  $t \geq 1$ . Suppose therefore that  $d_B(\xi_{t-1}) \geq \alpha 2^{-(t-1)}$ , where  $t \geq 1$ . To estimate  $d_B(\xi_t)$ , fix an arbitrary  $\xi \in B$ . Since  $B \subset \theta(B)$ , we can write  $\xi = \xi'/2 + j/2$  with  $\xi' \in B$  and  $j \in \{0, 1\}$ . If  $j \neq x_t^+$ , then  $|\xi_t - \xi| \geq \delta$ . (This is because  $B \subset (0, 1/2 - \delta) \cup (1/2 + \delta, 1)$ .) By the definition of  $\alpha$ , in this case we automatically have  $|\xi_t - \xi| \geq \alpha 2^{-t}$ . If  $x_t^+ = j$ , then

$$|\xi_t - \xi| = \left| \frac{\xi_{t-1}}{2} + \frac{x_t^+}{2} - \left( \frac{\xi'}{2} + \frac{j}{2} \right) \right| = |\xi_{t-1} - \xi'|/2 \geq d_B(\xi_{t-1})/2.$$

Using that  $d_B(\xi_{t-1}) \geq \alpha 2^{-(t-1)}$ , we again get  $|\xi_t - \xi| \geq \alpha 2^{-t}$ . Since  $\xi \in B$  was arbitrary,  $d_B(\xi_t) \geq \alpha 2^{-t}$ .

Now we prove (36). Let  $t \geq 1$ . To estimate  $d_{B_e^c}(\xi_t)$ , take  $\xi \in B_e^c$ . Then  $\theta(\xi) \in B$ , so we can write  $\xi = \xi'/2 + j/2$  with  $\xi' \in B$  and  $j \in \{0, 1\}$ . If  $j \neq x_t^+$ , we have  $|\xi_t - \xi| \geq \delta$ . (This follows from that  $B_e^c \subset (0, 1/2 - \delta) \cup (1/2 + \delta, 1)$ .) Thus  $|\xi_t - \xi| \geq \alpha 2^{-t}$  if  $j \neq x_t^+$ . If  $x_t^+ = j$ , then

$$|\xi_t - \xi| = \left| \frac{\xi_{t-1}}{2} + \frac{x_t^+}{2} - \left( \frac{\xi'}{2} + \frac{j}{2} \right) \right| = |\xi_{t-1} - \xi'|/2 \geq d_B(\xi_{t-1})/2,$$

so we again have  $|\xi_t - \xi| \geq \alpha 2^{-t}$ . Since  $\xi \in B_e^c$  was arbitrary,  $d_{B_e^c}(\xi_t) \geq \alpha 2^{-t}$ .  $\square$

*Proof of Theorem 1.* Let  $B \subset (0, 1)$  be closed and  $\theta$ -invariant with  $\overline{B_e^c} \cap \overline{B_b} = \emptyset$ .

**Step 1: Construction.** As in Proposition 2, take  $N_e$  and  $N_b$  to be closed disjoint subsets of  $[0, 1]$  with  $\overline{B_e^c} \subset N_e$  and  $\overline{B_b} \subset N_b$ , where  $N_e^* = N_b$ . Since  $\{0, 1/2, 1\} \cap$

$(B_e^c \cup B_b) = \emptyset$ , we may choose  $N_e$  and  $N_b$  so that  $\{0, 1/2, 1\} \cap (N_e \cup N_b) = \emptyset$ . Indeed, this can be achieved by taking the length of the intervals that define  $N_e$  in Proposition 2 to be sufficiently small. Finally, by increasing the length of these intervals if necessary, we can ensure that  $N_e$  contains  $\overline{B_e^c}$  in its interior. In summary,  $N_e$  and  $N_b$  are closed disjoint subsets of  $[0, 1]$  such that

- (i)  $\overline{B_e^c} \subset N_e$  and  $\overline{B_b} \subset N_b$ ,
- (ii)  $N_e^* = N_b$ ,
- (iii)  $\{0, 1/2, 1\} \cap (N_e \cup N_b) = \emptyset$ ,
- (iv)  $N_e$  contains  $\overline{B_e^c}$  in its interior.

Now choose  $\varepsilon > 0$  small so that  $N_e \cup N_b$  is disjoint from

$$N_0 := [0, \varepsilon] \cup [1 - \varepsilon, 1].$$

Then  $N_e \cup N_b$  is also disjoint from

$$N_{\frac{1}{2}} := [1/2 - \varepsilon, 1/2 + \varepsilon] = N_0^*.$$

Fix a positive integer  $k$  with  $2^{-k} < \varepsilon$ . Then  $k$  can be interpreted in the following way: starting from any  $\xi_0 \in [0, 1]$ , we can reach  $N_0$  by  $k$  consecutive steps to the left, or by  $k$  consecutive steps to the right. Set (we use the function (26))

$$p(\xi) = \begin{cases} |\log_2(d_{B_e^c}(\xi))|^{-1/k} & \text{if } \xi \in N_e \setminus \overline{B_e^c}, \\ 0 & \text{if } \xi \in \overline{B_e^c}, \\ 0 & \text{if } \xi \in N_{\frac{1}{2}}. \end{cases} \quad (37)$$

For  $\xi \in N_b \cup N_0$ , let  $p(\xi) = 1 - p(\xi^*)$ . Now  $p$  is defined on

$$N := N_e \cup N_b \cup N_{\frac{1}{2}} \cup N_0$$

and the equality in (30) holds on this set. Extend  $p$  to  $[0, 1/2]$  continuously and so that  $0 < p(\xi) < 1$  for all  $\xi \in [0, 1/2] \setminus (\overline{B_e^c} \cup N_{\frac{1}{2}} \cup \overline{B_b} \cup N_0)$ . Finally, if we set  $p(\xi) = 1 - p(\xi^*)$  for  $\xi \in [1/2, 1] \setminus N$  and extend periodically, we have a continuous  $p \in \mathcal{Q}$  with  $p(\xi) > 0$  for all  $\xi \in [0, 1] \setminus (\overline{B_e^c} \cup N_{\frac{1}{2}})$ . Since  $p \equiv 0$  on  $B_e^c$ ,  $B$  is  $p$ -invariant.

**Step 2: Conditions (a) and (b).** Because  $p$  takes the value one on  $[-\varepsilon, \varepsilon]$ , so does  $\hat{\Phi}_p$ , which means that (5) holds. To complete the proof we verify (6).

The basic conclusion in [6] says that

$$\sum_{k \in \mathbb{Z}} \hat{\Phi}_p(\xi_0 + k) = P(\xi_t \rightarrow 0 \text{ or } 1 | \xi_0) \quad (38)$$

for every  $\xi_0 \in [0, 1]$ . So to verify (6), it is enough to show that  $P(\xi_t \rightarrow 0 \text{ or } 1 | \xi_0)$  for Lebesgue almost every  $\xi_0 \in [0, 1]$ . Since  $B$  has measure zero (Lemma 1), we are done if we can show that  $P(\xi_t \rightarrow 0 \text{ or } 1 | \xi_0)$  for an arbitrary  $\xi_0 \in B^c$ .

Let  $\xi_0 \in B^c$  be arbitrary. That  $p \equiv 1$  on  $N_0$  means that if a sample path from  $\xi_0$  reaches  $N_0$ , it goes to 0 (if it reaches  $[0, \varepsilon]$ ) or 1 (if it reaches  $[1 - \varepsilon, 1]$ ). So it suffices to show that  $\xi_t \in N_0$  for some  $t$ ,  $P_p(\cdot | \xi_0)$ -a.s. By Levy's Borel-Cantelli Lemma (see [2] or [9, Lemma 4.1']), this is the case if

$$\sum_{n=0}^{\infty} P_p(\xi_{nk+k} \in N_0 | \xi_{nk}) = +\infty, \quad P_p(\cdot | \xi_0)\text{-a.s.} \quad (39)$$

We verify (39) by showing that there is a constant  $\lambda \in (0, 1)$  and  $a > 0$  such that

$$P_p(\xi_{nk+k} \in N_0 | \xi_{nk}) \geq \frac{\lambda}{a + nk + k} \quad (40)$$

for all  $n \geq 1$ ,  $P_p(\cdot | \xi_0)$ -a.s.

*Case 1:*  $\xi_{nk} \leq 1/2$ . If  $\xi_{nk} \leq 1/2$ , then  $\xi_{nk}/2^k \in N_0$  by our choice of  $k$ , so

$$\begin{aligned} P_p(\xi_{nk+k} \in N_0 | \xi_{nk}) &\geq P_p(\xi_{nk+k} = \xi_{nk}/2^k | \xi_{nk}) \\ &= \prod_{i=1}^k p(\xi_{nk}/2^i). \end{aligned}$$

That  $\xi_{nk} \leq 1/2$  implies  $\xi_{nk}/2^i \leq 1/4$  for all  $i \geq 1$ . So  $\xi_{nk}/2^i \in N_0 \cup N_e \cup N_b \cup N_r$  for  $i = 1, \dots, k$ , where

$$N_r := ([0, 1/4] \cap [3/4, 1]) \setminus (N_e \cup N_b \cup N_0).$$

The set of zeros of  $p$  is precisely  $\overline{B_e^c} \cup N_{\frac{1}{2}}$ . Since  $\overline{B_e^c}$  is in the interior of  $N_e$  and  $N_{\frac{1}{2}}$  is separated from  $N_r$  by a positive distance,  $\overline{N_r} \cap (\overline{B_e^c} \cup N_{\frac{1}{2}}) = \emptyset$ . Thus  $p(\xi) > 0$  for all  $\xi \in N_b \cup N_0 \cup \overline{N_r}$ . Since  $N_b \cup N_0 \cup \overline{N_r}$  is closed and  $p$  is continuous, we can choose  $c \in (0, 1)$  so that  $p$  is bounded below by  $c$  on  $N_b \cup N_0 \cup \overline{N_r}$ . All this means that  $\prod_{i=1}^k p(\xi_{nk}/2^i) \geq c^k$  if  $\xi_{nk}/2^i \in N_0 \cup N_b \cup N_r$  for  $i = 1, \dots, k$ . To achieve (40) we need a lower bound on  $\prod_{i=1}^k p(\xi_{nk}/2^i)$  for the case when  $\xi_{nk}/2^i \in N_e$  for at least one  $i \in \{1, \dots, k\}$ . By Lemma 2 we can choose  $\alpha_0 > 0$  so that  $d_{B_e^c}(\xi_t(x^+)) \geq 2^{-t-\alpha_0}$  for every sample path  $\xi_t(x^+)$  from  $\xi_0$ . Take  $x^+ \in \{0, 1\}^{\mathbb{N}}$  so that  $\xi_{nk+i}(x^+) = \xi_{nk}/2^i$ . (The first  $nk$  entries of  $x^+$  define the itinerary from  $\xi_0$  to  $\xi_{nk}$ , and  $x_{nk+i}^+ = 0$  for  $i = 1, \dots, k$ .) If  $\xi_{nk}/2^i \in N_e$ , the definition (37) of  $p$  together with (36) gives

$$\begin{aligned} p(\xi_{nk}/2^i) &= |\log_2(d_{B_e^c}(\xi_{nk}/2^i))|^{-1/k} = |\log_2(d_{B_e^c}(\xi_{nk+i}(x^+)))|^{-1/k} \\ &\geq |\log_2(2^{-(nk+i+\alpha_0)})|^{-1/k} \\ &= \left(\frac{1}{\alpha_0 + nk + i}\right)^{1/k}. \end{aligned}$$

Letting  $i_1, \dots, i_m$  be the  $m \in \{1, \dots, k\}$  integers  $i$  with  $\xi_{nk}/2^i \in N_e$ , we have

$$\prod_{i=1}^k p(\xi_{nk}/2^i) \geq c^{k-m} \prod_{j=1}^m \frac{1}{(\alpha_0 + nk + i_j)^{1/k}} \geq c^{k-m} \cdot \frac{1}{\alpha_0 + nk + k}.$$

This shows that (40) holds with  $a = \alpha_0$  and  $\lambda = c^k$ .

*Case 2:*  $\xi_{nk} > 1/2$ . If  $\xi_{nk} > 1/2$ , then  $N_0$  can be reached by  $k$  consecutive steps to the right:  $\xi_{nk+i} = \xi_{nk+i-1}/2 + 1/2$  for  $i = 1, \dots, k$ . We then have  $\xi_{nk+i} \geq 3/4$  and the above  $c$  bounds  $p(\xi_{nk+i})$  when  $\xi_{nk+i} \in N_b \cup N_0 \cup \overline{N_r}$ . Lemma 2 and the argument in Case 1 gives  $p(\xi_{nk+i}) \geq (\alpha_0 + nk + i)^{-1/k}$  when  $\xi_{nk+i} \in N_e$ . This means that (40) again holds with  $a = \alpha_0$  and  $\lambda = c^k$ .

Since  $n \geq 1$  was arbitrary, we have shown that (40) holds for all  $n \geq 1$  with  $a = \alpha_0$  and  $\lambda = c^k$ . We can now conclude that (39) holds.  $\square$

## 5. ON $g$ -FUNCTIONS FOR ONE-SIDED SUBSHIFTS

Let us say that a subshift  $K \subset X = \{0, \dots, J\}^{\mathbb{Z}^-}$  has a  $g$ -function if it is invariant for some continuous  $g \in \mathcal{G}$ , where  $\mathcal{G}$  denotes the  $g$ -functions on  $X$ , and that  $g \in \mathcal{G}$  is a *strict*  $g$ -function for  $K$  if  $g$  is continuous,  $K$  is  $g$ -invariant and  $g(\xi) > 0$  for



all  $\xi \in K$ . Before establishing necessary and sufficient conditions, we reemphasize that Krieger [14] has obtained sharp conditions for two-sided subshifts.<sup>4</sup> In fact the term “strict  $g$ -function” has been borrowed from [14]. Given a subshift  $S \subset X^\pm$ , Krieger defines

$$S_{(-\infty, n]} = \{(\dots, x_{n-1}, x_n) \in X : x^\pm \in S\}, n \geq 0,$$

$$\Gamma^+(x^-) = \{j \in \mathcal{A} : (x^-, j) \in S_{(-\infty, 1]}\}, x^- \in S_{(-\infty, 0]},$$

and says that  $S$  has a presentation by a  $g$ -function if there is a continuous function

$$G : \{(x^-, j) \in S_{(-\infty, 0]} \times \mathcal{A} : j \in \Gamma^+(x^-)\} \rightarrow [0, 1]$$

with the property that

$$\sum_{j \in \Gamma^+(x^-)} G(x^-, j) = 1 \text{ for every } x^- \in S_{(-\infty, 0]}.$$

If there exists such a function which is strictly positive, then  $S$  is said to have a presentation by a strict  $g$ -function.

The one and two-sided settings are equivalent in the sense that given  $\mathcal{F}$ ,  $X_{\mathcal{F}}$  has a (strict)  $g$ -function if and only if  $X_{\mathcal{F}}^\pm$  has a presentation by a (strict)  $g$ -function. Krieger shows that a subshift with a strict  $g$ -function presentation is of finite type, meaning  $S = X_{\mathcal{F}}^\pm$  for some finite  $\mathcal{F}$ . The conclusion is identical in the one-sided setting, where the result (Theorem 3) becomes a consequence of Proposition 1(c). By the equivalence, our condition

$$\bigcap_{j=0}^J (\overline{K_e^c})^{*,j} = \emptyset \quad (41)$$

for a  $g$ -function may be viewed as a translation of Krieger’s [14, p. 306-307] condition for a  $g$ -function presentation to the one-sided setting.

*Proof of Theorem 2.* If  $K$  is  $g$ -invariant, we must have  $g(\xi) = 0$  for all  $\xi \in \overline{K_e^c}$ . Since  $\bigcap_{j=0}^J (\overline{K_e^c})^{*,j} \subset \overline{K_e^c}$ , this means that  $\sum_{\xi' \in \theta^{-1}(\theta(\xi))} g(\xi') = 0$  if  $\xi \in \bigcap_{j=0}^J (\overline{K_e^c})^{*,j}$ . But then (1) fails, so (41) is necessary. For sufficiency we assume that (41) holds and construct a  $g \in \mathcal{G}$  with the property that  $\{\xi : g(\xi) = 0\} = \overline{K_e^c}$ . Choose  $m$  so that  $\xi \in \overline{K_e^c}, \xi' \in \bigcap_{j=1}^J (\overline{K_e^c})^{*,j}$  implies  $\rho(\xi, \xi') > 2^{-m}$  and let  $\bar{x} = \bar{x}_m$  be any word such that  $C(\bar{x}) \cap \overline{K_e^c}$  is nonempty. We begin by defining  $g$  on

$$U(\bar{x}) := \bigcup_{j=0}^J C(\bar{x}^{*,j}),$$

setting

$$g(\xi) = \begin{cases} \frac{1}{J} \rho(\xi, \overline{K_e^c}) & \text{if } \xi \in C(\bar{x}), \\ \frac{1}{J} (1 - \rho(\xi^{*, -j}, \overline{K_e^c})) & \text{if } \xi \in C(\bar{x}^{*,j}), j = 1, \dots, J. \end{cases}$$

Let  $\bar{x}'$  be another word of length  $m$  such that  $\overline{K_e^c} \cap C(\bar{x}') \neq \emptyset$ . If  $\bar{x}'^{*,j} = \bar{x}$  for some  $j$ , then  $C(\bar{x}') \subset U(\bar{x})$ ; if not,  $U(\bar{x})$  and  $U(\bar{x}') := \bigcup_{j=0}^J C(\bar{x}'^{*,j})$  are disjoint and we define  $g$  on  $U(\bar{x}')$  by the same recipe. This process is continued until  $g$  is defined on  $\bigcup_{j=0}^J \{\xi \in X : \rho(\xi, K_e^c) \leq 2^{-m}\}^{*,j}$ . The complement of this clopen set is invariant

<sup>4</sup>This was brought to the author’s attention during the process of writing this paper.

with respect to  $^{*,j}, j = 1, \dots, J$ . If we define  $g(\xi) = \frac{1}{J+1}$  on this set, we get a function  $g \in \mathcal{G}$  with  $\{\xi : g(\xi) = 0\} = \overline{K_e^c}$ .  $\square$

**Theorem 3.** *A subshift has a strict  $g$ -function if and only if it is of finite type.*

*Proof.* If  $K$  is not of finite type, then  $K \cap \overline{K_e^c}$  is nonempty by Proposition 1. Since any  $g$  with respect to which  $K$  is invariant must vanish on this set, there are no strict  $g$ -functions. If  $K$  is of finite type, then (41) is satisfied since each  $(K_e^c)^{*,j}$  is closed by Proposition 1(c) and the continuity of  $^{*,j}$ . With  $g$  as in the proof of Theorem 2, we have  $\{\xi : g(\xi) = 0\} = \overline{K_e^c}$ . Since  $K_e^c$  is closed,  $\overline{K_e^c} = K_e^c \subset K^c$ .  $\square$

**Remark 3.** *The proof (and statement) of Theorem 3 relies on properties of subshifts. The arguments in the proof of Theorem 2 on the other hand did use  $\Theta$ -invariance or the fact that  $K$  was closed. Consequently, the same proof establishes necessary and sufficient conditions for general subsets to be invariant with respect to continuous  $g \in \mathcal{G}$ . Given  $g \in \mathcal{G}$ , a set  $E \subset X$  is invariant (in the sense that processes that start in  $E$  remain in  $E$ ) if and only if  $g(\xi) = 0$  for all  $\xi \in E_e^c$ . In order for  $E \subset X$  to be invariant with respect to some (not necessarily continuous)  $g \in \mathcal{G}$ , we must have  $E \subset \Theta(E)$ . This condition is easily seen to be equivalent to  $\bigcap_{j=0}^J (E_e^c)^{*,j} = \emptyset$  by arguing as in Lemma 1(a) with  $\Theta$  in the role of  $\theta$ . A set  $E \subset X$  is invariant with respect to a continuous  $g \in \mathcal{G}$  if and only if  $\bigcap_{j=0}^J (\overline{E_e^c})^{*,j} = \emptyset$ .*

## 6. CONCLUDING REMARKS

To verify (38), one must “code”  $\mathbb{Z}$  into  $X^+$ . One such coding is given in [6]. A simplified coding is presented in [9, p. 67], where  $x^+(k), k \in \mathbb{Z}$ , is defined so that for every  $\xi_0 \in [0, 1]$  and  $t \geq 1$ ,

$$\xi_t(k) := \xi_t(x^+(k)) = \frac{\xi_0 + k}{2^t} \pmod{1}. \quad (42)$$

To recover it, recall that the recursion (12) that defines  $\xi_t(x^+)$  implies

$$\xi_t(x^+) = \frac{\xi_0 + \sum_{i=1}^t x_i 2^{i-1}}{2^t}, t \geq 1, \quad (43)$$

from which it follows that (42) holds for  $k \geq 0$  if  $x^+(k)$  is the binary representation. The code  $x^+(k)$  for a negative  $k$  is given by the unique sequence  $(x_1, x_2, \dots)$  with  $-k - 1 = \sum_{i=1}^{\infty} (1 - x_i) 2^{i-1}$ . One verifies that the coding  $k \mapsto x^+(k)$  achieves (42), so that

$$\prod_{j=1}^t p(\xi_j(k)) = P(\xi_t = \xi_t(k) | \xi_0) \quad (44)$$

for  $\xi_0 \in [0, 1]$  and  $k \in \mathbb{Z}$ . Letting  $t \uparrow \infty$  in (44) gives

$$\hat{\Phi}_p(\xi_0 + k) = P(x^+(k) | \xi_0). \quad (45)$$

By (43) we have that  $\xi_t(x^+)$  converges to 0 or 1 precisely on the set of  $x^+ \in X^+$  that terminate in either zeros or ones. Since this set coincides with  $\{x^+(k) : k \in \mathbb{Z}\}$ , (38) follows from (45) by summing over  $k$ .

If  $P(\xi_t \rightarrow 0 \text{ or } 1 | \xi_0) = 1$  for every  $\xi_0 \in [0, 1]$ , no closed subset of  $(0, 1)$  can be invariant. Another way to state this condition is to require a compact  $T \subset \mathbb{R}$

of Lebesgue measure one such that for each  $\xi \in [0, 1]$  we can find a  $k \in \mathbb{Z}$  with  $\xi + k \in T$ , where

$$\inf_{j \geq 1} \inf_{\xi \in T} p(\xi/2^j) > 0. \quad (46)$$

This is Cohen's [3] condition. To see that it rules out invariant proper subsets of  $(0, 1)$ , note that if  $p \in \mathcal{Q}$  and a set  $T$  with the stated properties is given, then for each  $\xi_0 \in [0, 1]$  we can choose  $k \in \mathbb{Z}$  such that  $\xi_0 + k \in T$ . From (42) and (46) we get  $p(\xi_j(k)) = p((\xi_0 + k)/2^j) > 0$ ,  $j \geq 1$ , which by (44) implies  $P(\xi_t = \xi_t(k) | \xi_0) > 0$  for all  $t \geq 1$ . Since  $\xi_0$  was arbitrary and  $\xi_t(k) \rightarrow 0$  (if  $k \geq 0$ ) or  $1$  (if  $k < 0$ ) as  $t \rightarrow \infty$ , no proper subset of  $(0, 1)$  can be invariant.

The discovery of continuous  $p$  for which  $\sum_{k \in \mathbb{Z}} \hat{\Phi}_p(\xi + k) = 1$  fails precisely on a set of measure zero was made in [6]. The term "inaccessible" set comes from [9], where the example from [6] is included in a class of such invariant sets obtained from subshifts of finite type. This paper describes their structure completely.

#### REFERENCES

- [1] Berbee, Henry. : Chains with infinite connections: uniqueness and Markov representation. *Prob. Th. and Rel. Fields.* **76**, (1987), 243–253.
- [2] Chen, L.H.Y. : A short note on the conditional Borel-Cantelli lemma. *Ann. Probab.* **6**, (1978), 699–700.
- [3] Cohen, A.: Ondelettes, analyses multirésolutions et filtres miroirs en quadrature. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **7**, (1990), 439–459.
- [4] Cohen, A. and Conze, J-P. : Régularité des bases d'ondelettes et mesures ergodiques. *Rev. Mat. Iberoamericana* **8**, (1992), 351–365.
- [5] Conze, J-P. and Raugi, A.: Fonctions harmoniques pour un opérateur de transition et applications. *Bull. Soc. Math. France* **118**, (1990), 273–310.
- [6] Dobrić, V. and Gundy, R. and Hitczenko, P.: Characterizations of orthonormal scale functions: a probabilistic approach. *J. Geom. Anal.* **10**, (2000), 417–434.
- [7] Doeblin, W. and Fortet, R.: Sur des chaînes à liaisons complètes. *Bull. Soc. Math. France.* **65**, (1937), 132–148.
- [8] Einsiedler, M. and Ward, T.: Ergodic theory with a view towards number theory, Graduate Texts in Mathematics 259, *Springer*, (2011), xviii+481.
- [9] Gundy, R.F.: Probability, ergodic theory, and low-pass filters. In *Topics in harmonic analysis and ergodic theory*, volume 444 of *Contemp. Math.*, (2007), 53–87.
- [10] Gundy, R.F.: Tilings, scaling functions, and a Markov process, *Notices Amer. Math. Soc.*, **57**, (2010), 1094–1104.
- [11] Gundy, R. F. and Jonsson, A. L.: Scaling functions on  $\mathbb{R}^2$  for dilations of determinant  $\pm 2$ . *Appl. Comput. Harmon. Anal.* **29**, (2010), 49–62.
- [12] Hernández, E. and Weiss, G.: A first course on wavelets. *Studies in Advanced Mathematics*. CRC Press, Boca Raton, FL, 1996. With a foreword by Yves Meyer.
- [13] Johansson, A. and Öberg, A.: Square summability of variations of  $g$ -functions and uniqueness of  $g$ -measures. *Math. Res. Lett.* **10**, (2003), 587–601.
- [14] Krieger, W.: On  $g$ -functions for subshifts. *IMS Lecture Notes-Monograph Series* **48**, (2006), 306–316.
- [15] Keane, M.: Strongly mixing  $g$ -measures. *Invent. Math.* **16**, (1972), 309–324.
- [16] Lawton, W.M.: Tight frames of compactly supported affine wavelets. *J. Math. Phys.* **31**, (1990), 1898–1901.
- [17] Lind, D. and Marcus, B.: An introduction to symbolic dynamics and coding. *Cambridge University Press*, Cambridge, 1995. xvi+495 pp.
- [18] Papadakis, M. and Šikić, H. and Weiss, G.: The characterization of low pass filters and some basic properties of wavelets, scaling functions and related concepts. *J. Fourier Anal. Appl.* **5**, (1999), 495–521.
- [19] The Wutam Consortium.: Basic properties of wavelets. *J. Fourier Anal. Appl.* **4**, (1998), 575–594.

- [20] Stenflo, Ö.: Uniqueness in  $g$ -measures. *Nonlinearity*. **16**, (2003), 403–410.
- [21] Stenflo, Ö.: Uniqueness of invariant measures for place-dependent random iterations of functions. In *Fractals in multimedia (Minneapolis, MN, 2001)*. volume 132 of *IMA Vol. Math. Appl.* (2002), 13–32. Springer, New York.

DEPARTMENT OF ENGINEERING SCIENCES AND MATHEMATICS, LULEÅ UNIVERSITY OF TECHNOLOGY, 97187 LULEÅ, SWEDEN

*E-mail address:* adam.jonsson@ltu.se